Biometrics 62, 1197–1206
December 2006

DOI: 10.1111/j.1541-0420.2006.00617.x

Effects of Residual Smoothing on the Posterior of the Fixed Effects in Disease-Mapping Models

Brian J. Reich,[1,*] James S. Hodges,[2] and Vesna Zadnik[3]

[1]Department of Statistics, North Carolina State University, 2501 Founders Drive, Box 8203, Raleigh, North Carolina 27695, U.S.A.

[2]Division of Biostatistics, School of Public Health, University of Minnesota, 2221 University Avenue SE, Suite 200, Minneapolis, Minnesota 55414, U.S.A.

[3]Epidemiology and Cancer Registries, Institute of Oncology Ljubljana, Zaloka 2, 1000 Ljubljana, Slovenia

[*]*email:* reich@stat.ncsu.edu

Summary. Disease-mapping models for areal data often have fixed effects to measure the effect of spatially varying covariates and random effects with a conditionally autoregressive (CAR) prior to account for spatial clustering. In such spatial regressions, the objective may be to estimate the fixed effects while accounting for the spatial correlation. But adding the CAR random effects can cause large changes in the posterior mean and variance of fixed effects compared to the nonspatial regression model. This article explores the impact of adding spatial random effects on fixed effect estimates and posterior variance. Diagnostics are proposed to measure posterior variance inflation from collinearity between the fixed effect covariates and the CAR random effects and to measure each region's influence on the change in the fixed effect's estimates by adding the CAR random effects. A new model that alleviates the collinearity between the fixed effect covariates and the CAR random effects is developed and extensions of these methods to point-referenced data models are discussed.

Key words: Collinearity; Conditional autoregressive prior; Diagnostics; Disease mapping; Spatial regression models.

1. Introduction

Spatially referenced public health data sets have become more available in recent years. A common objective when analyzing these data sets is to estimate the effect of covariates on region-specific disease rates while accounting for spatial correlation. As a motivating example, consider analyzing the relationship between socioeconomic factors and stomach cancer incidence in Slovenia for the years 1995–2001 using data originally presented by Zadnik and Reich (2006).

Figure 1a plots the $n = 192$ municipalities' observed standardized incidence ratio (SIR), that is, the ratio of observed (O_i) to expected (E_i, computed using indirect standardization) number of cases, where i indexes municipalities. Each region's socioeconomic status has been placed into one of five ordered categories by Slovenia's Institute of Macroeconomic Analysis and Development based on a composite score calculated from 1999 data. Figure 1b shows the centered version of this covariate (SEc). Both SIR and SEc exhibit strong spatial patterns. Western municipalities generally have low SIR and high SEc while eastern municipalities generally have high SIR and low SEc. This suggests a negative association between stomach cancer rates and socioeconomic status. A simple regression assuming the O_i are independent Poisson random variables with $\log\{E(O_i)\} = \log(E_i) + \alpha + \beta_{\mathrm{SEc}} * \mathrm{SEc}_i$ and flat priors on α and β_{SEc} (Model 1, Table 1) shows SEc is indeed related to SIR: under this nonspatial model β_{SEc} has posterior median -0.137 and 95% posterior interval (-0.17, -0.10).

To model these data correctly, the spatial structure must be incorporated into the analysis. A popular Bayesian disease-mapping model is the conditionally autoregressive (CAR) model of Besag, York, and Mollié (1991). Under this model, the observed numbers of cases in each region follow conditionally independent Poisson distributions with $\log\{E(O_i)\} = \log(E_i) + \beta_{\mathrm{SEc}} * \mathrm{SEc}_i + S_i + H_i$. This model has one fixed effect, β_{SEc}, which is given a flat prior. The intercept is not fixed but is the sum of two types of random effects: S captures spatial clustering and H captures region-wide heterogeneity. The H_i are modeled as independent draws from a normal distribution with mean zero and precision τ_h. Spatial dependence is introduced through the prior (or model) on $S = (S_1, \ldots, S_n)'$. The CAR model with L_2 norm (also called a Gaussian Markov random field) for S has improper density

$$p(S \mid \tau_s) \propto \tau_s^{(n-G)/2} \exp\left(-\frac{\tau_s}{2} S'QS\right), \quad (1)$$

where τ_s controls smoothing induced by this prior, larger values smoothing more than smaller; G is the number of "islands" (disconnected groups of regions) in the spatial structure (Hodges, Carlin, and Fan, 2003); and Q is $n \times n$ with nondiagonal entries $q_{ij} = -1$ if regions i and j are neighbors and 0 otherwise, and diagonal entries q_{ii} are equal to the

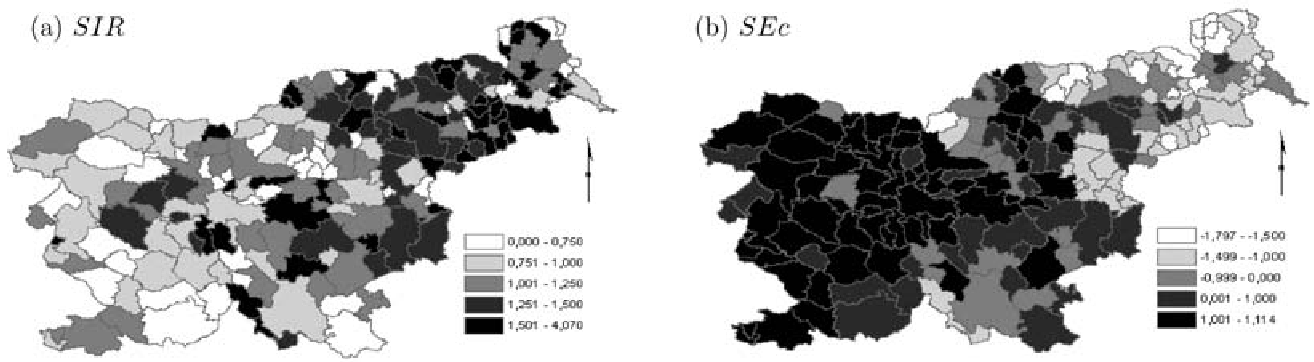

Figure 1. Each municipality's observed standardized incidence ratio (SIR = O/E) and centered socioeconomic status (SEc).

number of region i's neighbors. This is a multivariate normal kernel, specified by its precision matrix $\tau_s Q$ instead of the usual covariance matrix. In this article, we give the precision parameters τ_s and τ_h independent gamma(0.01, 0.01) priors parameterized to have mean 1 and variance 100.

Table 1 shows that Model 2 including the heterogeneity and spatial random effects has smaller deviance information criteria (DIC; Spiegelhalter et al., 2002) than the simple Poisson regression (1081.5 vs. 1153.0) despite the increase in model complexity (the effective number of parameters p_D increases from 2.0 to 62.3). Adding the random effects to the simple Poisson model also has a dramatic effect on β's posterior: its posterior mean changes from -0.137 to -0.022 and its posterior variance increases from 0.0004 to 0.0016.

This article proposes several diagnostics to investigate the change in the posterior of the fixed effects by adding spatial random effects. In the literature, diagnostics for the fixed effects in spatial regression have focused primarily on case deletion (e.g., Christensen, Johnson, and Pearson, 1992; Christensen, Pearson, and Johnson, 1993; Haslett, 1999). While measuring the change in the fixed effects' posterior due to removing individual observations is an important model-building step, it does not help explain the surprisingly large difference in β_{SEc}'s posterior under the spatial and nonspatial models. In this article, diagnostic measures analogous to leverage and the variance inflation factor for linear regression are developed to clarify the effect of adding spatial random effects on the posterior mean and variance of the fixed effects parameters.

The article proceeds as follows. Section 2 begins by analyzing the case of normally distributed outcomes, both for its own usefulness and because its more explicit theory sheds light on the Poisson case. We propose diagnostics to identify the effect of adding CAR random effects on the posterior mean and variance of the fixed effects parameters, and identify combinations of covariates and spatial grids that are especially troublesome. Section 3 presents one way to sidestep these collinearity problems by restricting the domain of the spatial random effects to the space orthogonal to the fixed effect covariates. The methods of Sections 2 and 3 are then extended to the generalized linear spatial regression model in Section 4. Section 5 illustrates these methods using the Slovenia stomach cancer data. Section 6 summarizes the results and briefly describes how the methods could be extended to other spatial models, such as geostatistical models for point-referenced data.

2. Collinearity in the CAR Model with Normal Observables

Let $\mathbf{y}$ be the n-vector of observed values and X be a known $n \times p$ matrix of covariates standardized so that each column sums to zero and has unit variance. The CAR spatial regression model is

$$\mathbf{y} \mid \boldsymbol{\beta}, S, \tau_e \sim N(X\boldsymbol{\beta} + S, \tau_e I_n) \tag{2}$$

$$S \mid \tau_s \sim N(0, \tau_s Q), \tag{3}$$

where $\boldsymbol{\beta}$ is a p-vector of fixed effect regression parameters, S is an n-vector, $\tau_e I_n$ and $\tau_s Q$ are precision matrices, and Q is the known adjacency matrix described in Section 1. Because the rows and columns of Q sum to zero, the CAR model necessarily implies a flat prior on each island's average of S. A common solution to this impropriety is to add fixed effects for

Table 1

DIC, $\bar{D}$, and p_D of various models along with 95% posterior confidence intervals for β (fixed effects parameters), τ_s (the prior precision of the CAR random effects), and τ_h (the prior precision of the heterogeneity random effects). The models are (1) simple Poisson regression, (2) spatial regression with CAR and heterogeneity random effects and a fixed effect for SEc, (3) Model 2 with a fixed effect for urban, and (4) Model 2 removing the CAR effects in the span of SEc.

Model	DIC	$\bar{D}$	p_D	β_{SEc}	β_{Urban}	τ_s	τ_h
1	1153.0	1151.0	2.0	(−0.175, −0.098)	—	—	—
2	1081.5	1019.2	62.3	(−0.100, 0.057)	—	(4.7, 77.4)	(21.0, 1224.0)
3	1081.8	1014.9	66.9	(−0.133, 0.044)	(−0.06, 0.25)	(5.4, 89.2)	(18.9, 248.1)
4	1088.0	1018.0	70.0	(−0.166, −0.069)	—	(4.9, 175.1)	(16.3, 236.8)

island's intercept and place a sum-to-zero constraint on each island's S. However, because collinearity between the intercept and the spatial random effects is not of interest, we let S remain unconstrained and assume X does not have a column for the intercept, so that the intercept is implicitly present in S. To complete a Bayesian specification, $\boldsymbol{\beta}$ is given a flat prior and τ_e and τ_s are given independent gamma(0.01, 0.01) priors.

An equivalent representation of (2)–(3) highlighting identification and collinearity concerns is

$$\mathbf{y} \,|\, \boldsymbol{\beta}, \mathbf{b}, \tau_e \sim N(X\boldsymbol{\beta} + Z\mathbf{b}, \tau_e I_n) \tag{4}$$

$$\mathbf{b} \,|\, \tau_s \sim N(0, \tau_s D), \tag{5}$$

where $\mathbf{b} = Z'S$ and $Q = ZDZ'$ represents Q's spectral decomposition for $n \times n$ orthogonal matrix Z and $n \times n$ diagonal matrix D with positive diagonal elements $d_1 \geq \cdots \geq d_{n-G}$. The last G diagonal elements of D are zero. The corresponding elements of $\mathbf{b}$ represent the intercepts of the G islands and are implicit fixed effects. The mean of $\mathbf{y}$ in (4) has $p + G$ fixed effects ($\boldsymbol{\beta}$, $b_{n-G+1}, \ldots, b_n$) and $n - G$ random effects ($b_1, \ldots, b_{n-G}$) for a total of $n + p$ predictors; given that there are only n observations, this raises identifiability and collinearity concerns. Each column of X is a linear combination of the n orthogonal columns of Z. Therefore, ignoring $\mathbf{b}$'s prior, that is, setting $\tau_s = 0$, the data cannot identify $\boldsymbol{\beta}$ and $\mathbf{b}$ in the sense that $\boldsymbol{\beta}$'s marginal posterior is proportional to its prior, that is, $p(\boldsymbol{\beta} \,|\, \mathbf{y}) \propto p(\boldsymbol{\beta})$. Identification of $\boldsymbol{\beta}$ relies on smoothing of $\mathbf{b}$, which is controlled by τ_s. As τ_s increases, $\mathbf{b}$ is smoothed closer to zero and the posterior of $\boldsymbol{\beta}$ becomes more similar to its posterior under the ordinary linear model (OLM). This effect of τ_s on the fixed effect's posterior is illustrated in the analysis of the Slovenian data in Section 5.

To measure the influence on the posterior mean and variance of $\boldsymbol{\beta}$ from including and smoothing the spatial random effects, we investigate the posterior of $\boldsymbol{\beta}$ conditional on (τ_e, τ_s). After integrating out $\mathbf{b}$, $\boldsymbol{\beta}$ is normal with

$$\begin{aligned} E(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y}) &= E\{E(\boldsymbol{\beta} \,|\, \mathbf{b}, \tau_s, \tau_e, \mathbf{y})\} \\ &= E\{(X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b}) \,|\, \tau_s, \tau_e, \mathbf{y}\} \\ &= (X'X)^{-1}X'(\mathbf{y} - Z\hat{\mathbf{b}}) \\ &= \hat{\boldsymbol{\beta}}_{\text{OLM}} - (X'X)^{-1}X'Z\hat{\mathbf{b}} \end{aligned} \tag{6}$$

$$\begin{aligned} \text{var}^{-1}(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y}) &= \tau_e X'X - X'\text{var}(Z\mathbf{b} \,|\, \boldsymbol{\beta}, \tau_e, \tau_s, \mathbf{y})X \\ &= \tau_e X'Z\tilde{D}Z'X, \end{aligned} \tag{7}$$

where $\hat{\boldsymbol{\beta}}_{\text{OLM}} = (X'X)^{-1}X'\mathbf{y}, \hat{\mathbf{b}} = E(\mathbf{b} \,|\, \tau_e, \tau_s, \mathbf{y}) = (Z'P^cZ + rD)^{-1}Z'P^c\mathbf{y}$ is $\mathbf{b}$'s posterior mean given (τ_s, τ_e) but not $\boldsymbol{\beta}$, var($\mathbf{b} \,|\, \boldsymbol{\beta}$, τ_e, τ_s, $\mathbf{y}$) $= (I + rQ)^{-1}$ is $\mathbf{b}$'s full conditional variance, $P^c = I - X(X'X)^{-1}X'$, $r = \tau_s/\tau_e$, and $\tilde{D} = I - (I + rD)^{-1}$. The posterior mean and variance of $\boldsymbol{\beta}$ under the OLM with $\mathbf{b} \equiv 0$ are $\hat{\boldsymbol{\beta}}_{\text{OLM}}$ and $(\tau_e X'X)^{-1}$, respectively. By adding the random effects, $\boldsymbol{\beta}$'s posterior mean is shifted by $(X'X)^{-1}X'Z\hat{\mathbf{b}}$, (6), and $\boldsymbol{\beta}$'s posterior precision (inverse variance) is decreased by $X'\text{var}(\mathbf{b} \,|\, \boldsymbol{\beta}, \tau_e, \tau_s)X$, (7). These two effects are discussed separately in Sections 2.1 and 2.2.

2.1 *Influence of Spatial Random Effects on* $E(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y})$

This section provides diagnostics measuring each region's contribution to the difference between $E(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y})$ and $\hat{\boldsymbol{\beta}}_{\text{OLM}}$. Case influence on the estimates in the OLM has a long history, with two common diagnostics being leverage and the DFBETA statistics (Hocking, 1996). For the OLM, the leverage of the ith observation is the ith diagonal element of $X(X'X)^{-1}X'$ and measures the ith observation's potential to be overly influential in fixed effect estimation. The leverages are properties of the design and do not depend on the outcomes y_i; observations with outlying X_i typically have large leverage. By contrast, the DFBETA statistics are functions of the observed data, $\mathbf{y}$. They measure the influence of the ith observation on the estimate of $\boldsymbol{\beta}_j$ by comparing the estimate of $\boldsymbol{\beta}_j$ with and without the ith observation included in the analysis. This section develops measures of potential and actual influence analogous to leverage and DFBETA diagnostics.

The posterior mean of $\boldsymbol{\beta}$ given (τ_e, τ_s) under the spatial model is $(X'X)^{-1}X'(y - X\hat{\mathbf{b}})$ (equation (6)), the least squares estimate using the residuals $\mathbf{y} - Z\hat{\mathbf{b}}$ as the observables. Define the change in $\boldsymbol{\beta}$'s posterior mean given (τ_e, τ_s) due to adding the CAR random effects as $\Delta = \hat{\boldsymbol{\beta}}_{\text{OLM}} - E(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y}) = (X'X)^{-1}X'Z\hat{\mathbf{b}}$. Δ can be positive or negative, that is, adding the spatial random effects does not necessarily push the posterior mean of $\boldsymbol{\beta}$ toward zero as in the example in Section 1.

Δ can be written as $\Delta = C\mathbf{e}$, where C is the $p \times n$ matrix,

$$\begin{aligned} C &= \{(X'X)^{-1}X'\}\{Z(Z'P^cZ + rD)^{-1}Z'\}P^c \\ &= \{(X'X)^{-1}X'\}\{(P^c + rQ)^{-1}\}P^c, \end{aligned} \tag{8}$$

and $\mathbf{e} = P^c\mathbf{y}$ is the n-vector of OLM residuals. As for leverage, the elements of C depend on the design (i.e., the spatial structure and the covariates) but not on the data except through $r = \tau_s/\tau_e$, on which we have conditioned. Plotting the absolute values of the jth row of C, which we call $|C_j|$, shows which regions have the potential to make a large contribution to the change in β_j when spatial effects are added to the OLM.

Section 4 develops an approximation for C that is more appropriate for analyzing Poisson outcomes, as in Section 1's Slovenian data. Nonetheless, we analyze (8) here for the Slovenia design to compare (8) with well-understood measures of influence for the OLM. Figure 2a plots $|C|$ for the Slovenia spatial grid with one covariate, $X = $ SEc, and $r = 0.1$. A careful comparison of Figure 1b and Figure 2a shows that regions with extreme SEc (i.e., high leverage in OLM) do not necessarily have large $|C_i|$; the regions with large $|C_i|$ are those with extreme SEc relative to the SEc of nearby regions. For example, most regions with large SEc are in western Slovenia (Figure 1b). However, only western municipalities with moderate SEc have large $|C_i|$ because moderate SEc is unusual relative to other western regions.

Figure 2b plots $|C|$ with little spatial smoothing ($r = 0.1$) on the horizontal axis and substantial spatial smoothing ($r = 10$) on the vertical axis. $|C|$ is generally smaller with $r = 10$ because for large r the spatial random effects are smoothed to zero and β's posterior is similar to its posterior under the OLM. Also, the number of regions that qualify as "nearby" is determined by r. Consider the northwestern municipality Jesenice indicated in Figure 2. For small r (Figure 2a), there

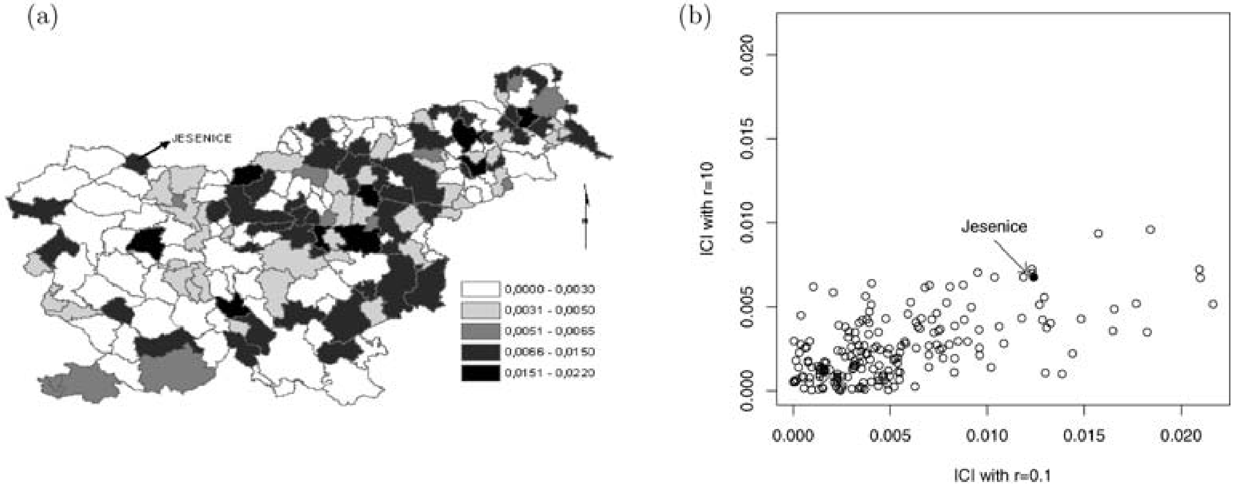

Figure 2. Plot of $|C|$ for Slovenian municipalities with $X =$ SEc. (a) shows $|C|$ with $r = 0.1$ and (b) compares $|C|$ with $r = 0.1$ and $r = 10$.

is little spatial smoothing and Jesenice's SEc is primarily compared to its immediately neighboring regions. Its SEc is smaller than the SEc of the three adjacent municipalities (Figure 1b), so its $|C_i|$ is moderately large. However, for $r = 10$ (Figure 2b, vertical axis) Jesenice's $|C_i|$ is one of the largest because for large r (strong spatial smoothing) its SEc is compared not only to adjacent regions, but to all northwestern regions. Because Jesenice has the smallest SEc in the entire northwest and few other regions are extreme relative to such a large group of nearby regions, its $|C_i|$ is now among the largest.

Only regions with nonzero C_{ij} and nonzero e_i contribute to $\Delta_j = \sum_{i=1}^n \delta_{ij} = \sum_{i=1}^n C_{ij} e_i$. In the OLM, $\hat{\boldsymbol{\beta}}_j$ is computed assuming all associations between nearby regions can be explained by the fixed effects. C_{ij} measures the ith region's potential to update this estimate of β_j when spatial clustering is introduced as a second explanation of association between nearby regions. Sites with covariates that are similar to neighboring sites (small $|C_{ij}|$) cannot distinguish between these competing explanations of association because both spatial clustering and the fixed effects predict the region will be similar to its neighbors. Regions with extreme covariates relative to nearby regions (large $|C_{ij}|$) can provide information to update $\hat{\boldsymbol{\beta}}_j$ because the two explanations of association predict different outcomes. The OLM predicts that y_i will be different from its neighbors. If y_i really is different from its neighbors (small $|e_i|$) this supports the assumption that all associations are caused by fixed effects and an update to $\hat{\boldsymbol{\beta}}_j$ is not warranted. In contrast, if y_i is similar to the outcomes of nearby regions (large $|e_i|$), this is evidence that spatial clustering, not fixed effects, is responsible for the similarity of nearby regions. This suggests a large change in $\boldsymbol{\beta}_j$, that is, $|\delta_{ij}|$ will be large. That is, if the simple linear regression fits poorly at regions where X_{ij} is different from neighboring regions' X_{ij}, there will be a large adjustment to the posterior mean of $\boldsymbol{\beta}_j$ when the spatial structure is incorporated into the model.

2.2 *Effect of Collinearity on var*$(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y})$ *and var*$(\boldsymbol{\beta} \,|\, r, \mathbf{y})$

The variance inflation factor is commonly used in OLM theory to measure collinearity among covariates (Hocking, 1996). This diagnostic is defined as the actual variance of the jth regression parameter divided by its variance assuming the jth covariate is uncorrelated with the other covariates. Although collinearity among fixed effects is an important issue, in this section we will inspect the increase in $\boldsymbol{\beta}_j$'s posterior variance by adding CAR random effects to the OLM.

Because the columns of Z are an orthogonal basis for $\Re^n$, we can write $X = (n - 1Z\boldsymbol{\rho})^{\frac{1}{2}}$ for an $n \times p$ matrix $\boldsymbol{\rho}$ where ρ_{ij} is the correlation between the ith column of Z and the jth column of X and $\sum_{i=1}^n \rho_{ij}^2 = \text{var}(X_j) = 1$. The posterior variance of $\boldsymbol{\beta}$ in (7) can be written $\text{var}(\boldsymbol{\beta} \,|\, \tau_e, \tau_s, \mathbf{y}) = \{(n-1)\tau_e \boldsymbol{\rho}' \tilde{D} \boldsymbol{\rho}\}^{-1}$, where $\tilde{D} = I - (I + rD)^{-1} = \text{diag}\{rd_i/(1 + rd_i)\}$. Under the OLM with $\tau_s = \infty$, $\text{var}(\boldsymbol{\beta} \,|\, \tau_e, \mathbf{y}) = (\tau_e X'X)^{-1} = \{(n-1)\tau_e \boldsymbol{\rho}'\boldsymbol{\rho}\}^{-1}$, so

$$\text{VIF}_j(\boldsymbol{\beta} \,|\, r, \tau_e) = \frac{(X'Z\tilde{D}Z'X)^{-1}_{jj}}{(X'X)^{-1}_{jj}} = \frac{(\boldsymbol{\rho}'\tilde{D}\boldsymbol{\rho})^{-1}_{jj}}{(\boldsymbol{\rho}'\boldsymbol{\rho})^{-1}_{jj}} \tag{9}$$

is the inflation in $\boldsymbol{\beta}_j$'s variance by adding the CAR random effects. Although we have conditioned on τ_e, $\text{VIF}_j(\boldsymbol{\beta} \,|\, r, \tau_e)$ depends only on $r = \tau_s/\tau_e$, not τ_e. The relationship between τ_e and variance inflation is discussed below.

If X is a vector, that is, the number of covariates is $p = 1$, (9) reduces to

$$\text{VIF}(\beta \,|\, r, \tau_e) = \left(1 - \sum_{i=1}^n \frac{\rho_i^2}{1 + rd_i}\right)^{-1} \geq 1 \text{ for all } r > 0. \tag{10}$$

For any spatial grid and for any covariate, conditional on τ_e, $\boldsymbol{\beta}$'s posterior variance is larger under the spatial model than the OLM for any $r > 0$. $\text{VIF}(\beta \,|\, r, \tau_e) \to \infty$ as $r \to 0^+$, that is, $\boldsymbol{\beta}$ has infinite variance if the random effects $\mathbf{b}$ are not smoothed. As r increases, $\mathbf{b}$ is smoothed to zero and $\text{VIF}(\beta \,|\, r, \tau_e) \to 1$. The rate at which $\text{VIF}(\beta \,|\, r, \tau_e)$ descends to 1 depends on the ρ_i and the corresponding d_i. $\text{VIF}(\beta \,|\, r, \tau_e)$ approaches 1 less quickly if the eigenvectors Z_i with large ρ_i are associated with small d_i. Reich and Hodges (2005) show that the eigenvectors associated with small eigenvalues d_i are generally "low-frequency" eigenvectors, that is, those that vary gradually in space. For example, the eigenvector with the smallest eigenvalue for the Slovenian grid describes the southwest/northeast gradient. In summary, variance inflation arising from collinearity between fixed effects and random effects is most troublesome when r is small and/or the covariates are highly correlated with low-frequency eigenvectors of the spatial adjacency matrix Q.

If $\rho_i \approx 1$ for some i, the vector of covariates X is very similar to the ith column of Z, so $\boldsymbol{\beta}$ and $\mathbf{b}_i$ must compete to explain the same one-dimensional projection of y (4). Both parameters are identified if $r > 0$ because $\mathbf{b}_i$'s prior shrinks $\mathbf{b}_i$ toward its prior mean, zero (5). If $\mathbf{b}_i$'s prior is strong, $\boldsymbol{\beta}$ will be well identified; if $\mathbf{b}_i$'s prior is weak, $\boldsymbol{\beta}$ will be weakly identified. The strength of $\mathbf{b}_i$'s prior relative to the prior on the other elements of $\mathbf{b}$ is controlled by d_i. The inflation of $\boldsymbol{\beta}$'s variance due to collinearity with Z_i persists for larger r if X is highly correlated with Z_i corresponding to small d_i because small d_i means $\mathbf{b}_i$'s prior is less restrictive.

Of course, when τ_e is unknown its estimate will generally be different under the spatial model from the OLM, which can indirectly affect β's variance inflation. In practice, it might be useful to investigate β's variance after integrating out both $\boldsymbol{\theta}$ and τ_e. If X is a vector (i.e., $p = 1$) and τ_e is given a gamma($a_e/2$, $b_e/2$) prior under both models, the ratio of β's variance under the spatial and nonspatial models after integrating out ($\boldsymbol{\theta}$, τ_e) is

$$\mathrm{VIF}(\beta \mid r) = \frac{B(r) + b_e}{n - 3 - G + a_e}\left(\frac{B(\infty) + b_e}{n - 3 + a_e}\right)^{-1}\mathrm{VIF}(\beta \mid r, \tau_e), \tag{11}$$

where $\mathrm{VIF}(\beta \mid r,\ \tau_e)$ is the variance inflation factor conditional on r and τ_e in (10), $B(\infty) = \mathbf{y}'P^c\mathbf{y}$, $B(r) = \mathbf{y}'(\Gamma - \Gamma X(X'\Gamma X)^{-1}X'\Gamma)\mathbf{y}$, $\Gamma = I - (I + rQ)^{-1}$, and G is the number of spatial islands. A plot of $\mathrm{VIF}(\beta \mid r)$ over a range of r could be used to illustrate the variance inflation of β due to adding the spatial random effects and might suggest a prior for r, analogous to ridge regression, that ensures variance inflation will be less than a particular value by prohibiting small values of r, that is, requiring at least a certain amount of smoothing.

3. Spatial Smoothing Orthogonal to the Fixed Effects

Several remedial measures have been proposed for collinearity in the OLM including variable deletion, principal component regression, and ridge regression (Hocking, 1996). If a pair of covariates is highly correlated, the natural reaction is variable deletion, that is, remove one of the covariates from the model. The estimates of the remaining coefficients will be more precise but potentially biased. In situations where both correlated covariates are of scientific interest, it may be difficult to decide which covariate to remove, and removing a scientifically relevant covariate may result in a model that is difficult to interpret.

Collinearity between the fixed effects and CAR random effects in Section 2's spatial regression model leads to many of the same problems as collinearity in the OLM, such as variance inflation and computational problems. However, a natural ordering of parameter importance may distinguish spatial regression from the OLM. The primary objective in a spatial regression may be to estimate the fixed effects; in such a situation, the CAR random effects are added merely to account for spatial correlation in the residuals when computing the posterior variance of the fixed effects, or to improve predictions. With these objectives in mind, it may be reasonable to proceed by removing the combinations of CAR random effects that are collinear with the fixed effects.

The orthogonal projection matrices $P = X(X'X)^{-1}X'$ and $P^c = I - P$ have ranks p and $n - p$, respectively. Let K be the $p \times n$ matrix whose rows are the p eigenvectors of P corresponding to nonzero eigenvalues. Similarly, let L be the $(n - p) \times n$ matrix whose rows are the $n - p$ eigenvectors of P^c corresponding to nonzero eigenvalues. $\boldsymbol{\theta}_1 = KZb$ are the combinations of the CAR random effects in the span of X, while $\boldsymbol{\theta}_2 = LZb$ are the combinations of the CAR random effects that are orthogonal to X. Transforming from $\mathbf{b}$ to $\boldsymbol{\theta} = (\boldsymbol{\theta}_1',\ \boldsymbol{\theta}_2')'$ gives the model

$$\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \tau_e \sim N(X\boldsymbol{\beta} + K'\boldsymbol{\theta}_1 + L'\boldsymbol{\theta}_2, \tau_e I_n) \tag{12}$$

$$\boldsymbol{\theta} \mid \tau_s \sim N(0, \tau_s\tilde{Q}), \tag{13}$$

where $\tilde{Q} = (K'L')'Q(K'L')$ and $\tau_e I$ and $\tau_s\tilde{Q}$ are precision matrices.

Identifying both $\boldsymbol{\beta}$ and $\boldsymbol{\theta}_1$ is entirely dependent on prior information because X and K' have the same span. Because $\boldsymbol{\beta}$ is given a flat prior, it is free to explain all $\mathbf{y}$'s variation in the span of X or K', so it is not clear whether the data identify $\boldsymbol{\theta}_1$ at all. After integrating $\boldsymbol{\beta}$ out of (12) and (13), the posterior of ($\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$) given (τ_e, τ_s) can be written as

$$p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \mid \tau_e, \tau_s, \mathbf{y}) \propto \tau_e^{(n-p)/2}\exp\left\{-\frac{\tau_e}{2}(L\mathbf{y} - \boldsymbol{\theta}_2)'(L\mathbf{y} - \boldsymbol{\theta}_2)\right\} \tag{14}$$

$$\times\ \tau_s^{(n-G)/2}\exp\left\{-\frac{\tau_s}{2}\begin{pmatrix}\boldsymbol{\theta}_1\\ \boldsymbol{\theta}_2\end{pmatrix}'\begin{pmatrix}\tilde{Q}_{11} & \tilde{Q}_{12}\\ \tilde{Q}_{12}' & \tilde{Q}_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{\theta}_1\\ \boldsymbol{\theta}_2\end{pmatrix}\right\}, \tag{15}$$

where $\tilde{Q}_{11} = KQK'$, $\tilde{Q}_{12} = KQL'$, and $\tilde{Q}_{22} = LQL'$. The likelihood (14) involves only $L\mathbf{y}$ and $\boldsymbol{\theta}_2$; neither $\boldsymbol{\theta}_1$ nor the part of $\mathbf{y}$ in the span of $X(K\mathbf{y})$ remain in the likelihood after integrating out the fixed effects. Assuming $\tilde{Q}_{12} \neq 0$, the data indirectly identify $\boldsymbol{\theta}_1$ through $\boldsymbol{\theta}_1$'s prior correlation with $\boldsymbol{\theta}_2$ (15), so the strength of $\boldsymbol{\theta}_1$'s identification depends on τ_s and $\tilde{Q}_{12}$. If τ_s is small or $\tilde{Q}_{12} \approx 0$, the identification of $\boldsymbol{\theta}_1$ is weak. If each column of X is an eigenvector of $Q, \tilde{Q}_{12} = 0$, and $\boldsymbol{\theta}_1$ is not identifiable.

As with variable deletion in the ordinary linear model, setting $\boldsymbol{\theta}_1$ to zero alleviates collinearity. With $\boldsymbol{\theta}_1 \equiv 0$, the model in (12) and (13) becomes

$$\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\theta}_2, \tau_e \sim N(X\boldsymbol{\beta} + L'\boldsymbol{\theta}_2, \tau_e I_n) \tag{16}$$

$$\boldsymbol{\theta}_2 \mid \tau_s \sim N(0, \tau_s\tilde{Q}_{22}). \tag{17}$$

Conditional on (τ_e, τ_s), $\boldsymbol{\beta}$'s posterior under this model is the same as its posterior under the OLM, normal with mean $\hat{\boldsymbol{\beta}}_{\mathrm{OLM}}$ and variance $(\tau_e X'X)^{-1}$. In spatial regression, two factors can inflate β_{SEc}'s posterior variance compared to the OLM: collinearity with the spatial random effects and reduction in the effective number of observations because of spatial correlation in the data. The reduced model in (16) and (17) removes any increase due to collinearity, but the marginal posterior of $\boldsymbol{\beta}$ will have larger variance than the marginal posterior of $\boldsymbol{\beta}$ from OLM because τ_e is stochastically decreased by the spatial correlation of the residuals.

To see this, first transform from (τ_e, τ_s) to $(\tau_e, r = \tau_s/\tau_e)$. After integrating out $\boldsymbol{\beta}$ and $\mathbf{b}$, the marginal posterior of τ_e given r is

$$\tau_e \mid r, \mathbf{y} \sim \text{gamma}\Big[0.5(n-p) + a_e,\ 0.5(L\mathbf{y})' \times \Big\{I - (I + r\tilde{Q}_{22})^{-1}\Big\}(L\mathbf{y}) + b_e\Big], \tag{18}$$

where τ_e's prior is gamma$(a_e,\ b_e)$. The marginal posterior of τ_e from OLM is gamma$(0.5(n-p) + a_e, 0.5(L\mathbf{y})'(L\mathbf{y}) + b_e)$. Because $(L\mathbf{y})'\{I - (I + r\tilde{Q}_{22})^{-1}\}(L\mathbf{y}) < (L\mathbf{y})'(L\mathbf{y})$ for all r, the marginal posterior of the error precision in the reduced CAR model in (16) and (17) is stochastically smaller than the marginal posterior of the error precision in the OLM.

4. Influence and Collinearity in the CAR Model with Nonnormal Observables

As is often the case with disease-mapping models, Section 1's analysis of the Slovenia stomach cancer data assumed that outcomes followed Poisson distributions. Spatial models for non-Gaussian data were popularized by Clayton and Kaldor (1987). Although the intuition gained from studying collinearity in the linear model with Gaussian errors is useful in studying collinearity in the generalized linear model, several authors have shown that the diagnostics used in linear models can be misleading when applied to generalized linear models (Schaefer, 1979; Mackinnon and Puterman, 1998). Therefore, in this section we extend the methods of Sections 2 and 3 to the generalized linear spatial regression model.

Assume the observations y_i are independent given η_i with conditional log likelihood

$$\log\{p(y_i \mid \eta_i)\} = y_i\eta_i - b(\eta_i) + c(y_i), \tag{19}$$

where $E(y_i \mid \eta_i) = \mu_i$, and $g(\mu_i) = \eta_i = X_i\boldsymbol{\beta} + S_i + H_i$ for some link function g. The linear predictor η is the sum of three vectors: $X\boldsymbol{\beta}$, where X is a known $n \times p$ matrix of covariates and $\boldsymbol{\beta}$ is a p-vector of fixed effects; the spatial random effects S, which follow a CAR$(\tau_s Q)$ distribution; and heterogeneity random effects H, which follow a normal$(0, \tau_h I_n)$ distribution, where $\tau_h I_n$ is a precision matrix.

Following Section 2, the linear predictor can be rewritten to highlight collinearity issues. After reparameterizing from S to $\mathbf{b} = Z'S$ where $Q = ZDZ'$ for $n \times n$ orthogonal matrix Z and $n \times n$ diagonal matrix D with positive diagonal elements $d_1 \geq \cdots \geq d_{n-G}$, the linear predictor is $\eta = X\boldsymbol{\beta} + Z\mathbf{b} + H$ where $\mathbf{b}$'s prior is normal with mean zero and precision $\tau_s D$.

The diagnostics of Sections 2.1 and 2.2 required closed-form expressions of $\boldsymbol{\beta}$'s posterior mean and variance. Non-Gaussian likelihoods prohibit closed-form expressions but these diagnostics can be extended to the non-Gaussian case using approximate methods (Lee and Nelder, 1996). One method for computing the posterior mode of $p(\boldsymbol{\beta},\ \mathbf{b},\ H \mid \tau_h,\ \tau_s,\ \mathbf{y})$ is iteratively reweighted least squares (IRLS). This method begins with an initial value, $(\boldsymbol{\beta}^{(1)},\ \mathbf{b}^{(1)},\ H^{(1)})$, and the new value at $(t+1)$th iteration is computed using the recurrence equation

$$\begin{pmatrix} \boldsymbol{\beta}^{(t+1)} \\ \mathbf{b}^{(t+1)} \\ H^{(t+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}^{(t)} \\ \mathbf{b}^{(t)} \\ H^{(t)} \end{pmatrix} - h(\boldsymbol{\beta}^{(t)}, \mathbf{b}^{(t)}, H^{(t)})^{-1} \times \begin{pmatrix} X' \\ Z' \\ I \end{pmatrix} \Big\{y - \mu\Big(\boldsymbol{\beta}^{(t)}, \mathbf{b}^{(t)}, H^{(t)}\Big)\Big\}, \tag{20}$$

where $\mu(\boldsymbol{\beta},\ \mathbf{b},\ H)_i = E(y_i \mid \boldsymbol{\beta},\ \mathbf{b},\ H)$,

$$h(\boldsymbol{\beta}, \mathbf{b}, H) = \begin{pmatrix} X'W(\boldsymbol{\beta}, \mathbf{b}, H)X & X'W(\boldsymbol{\beta}, \mathbf{b}, H)Z & X'W(\boldsymbol{\beta}, \mathbf{b}, H) \\ Z'W(\boldsymbol{\beta}, \mathbf{b}, H)X & Z'W(\boldsymbol{\beta}, \mathbf{b}, H)Z + \tau_s D & Z'W(\boldsymbol{\beta}, \mathbf{b}, H) \\ W(\boldsymbol{\beta}, \mathbf{b}, H)X & W(\boldsymbol{\beta}, \mathbf{b}, H)Z & W(\boldsymbol{\beta}, \mathbf{b}, H) + \tau_h I \end{pmatrix} \tag{21}$$

is the Hessian matrix, and $W(\boldsymbol{\beta},\ \mathbf{b},\ H)$ is diagonal with $W(\boldsymbol{\beta},\ \mathbf{b},\ H)_{ii} = \text{var}(y_i \mid \boldsymbol{\beta},\ \mathbf{b},\ H)$.

To assess each region's contribution to the difference between the posterior mode of $p(\boldsymbol{\beta} \mid \tau_e,\ \tau_s,\ \mathbf{y})$ with $(\hat{\boldsymbol{\beta}})$ and without $(\boldsymbol{\beta}^*)$ the random effects in the model, we approximate $\hat{\boldsymbol{\beta}}$ using one IRLS step with initial value $(\boldsymbol{\beta},\ \mathbf{b},\ H) = (\boldsymbol{\beta}^*,\ 0,\ 0)$, that is,

$$\begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \\ \hat{H} \end{pmatrix} \approx \begin{pmatrix} \boldsymbol{\beta}^* \\ 0 \\ 0 \end{pmatrix} - h(\boldsymbol{\beta}^*, 0, 0)^{-1} \begin{pmatrix} X' \\ Z' \\ I \end{pmatrix} \{y - \mu(\boldsymbol{\beta}^*, 0, 0)\}. \tag{22}$$

Using this approximation, $\Delta = \boldsymbol{\beta}^* - \hat{\boldsymbol{\beta}} \approx C\{\mathbf{y} - \mu(\boldsymbol{\beta}^*, 0, 0)\}$ where C is the $p \times n$ matrix whose rows are the first p rows of $h(\boldsymbol{\beta}^*,\ 0,\ 0)^{-1}(X\ \ Z\ \ I)'$. As with the diagnostic for normal data in Section 2.1, this approximation of Δ is the product of a $p \times n$ matrix and the n-vector of residuals from the fit without spatial random effects. In Section 5 the $\delta_{ij} = C_{ij}\{y_i - \mu(\boldsymbol{\beta}^*,\ 0,\ 0)_i\}$ are plotted to search for municipalities of Slovenia that are highly influential in the change of β_{SEc} from the nonspatial to the spatial regression model.

Using Fisher's observed information, the posterior variance of $(\boldsymbol{\beta},\ \mathbf{b},\ H)$ given $(\tau_h,\ \tau_s)$ can be approximated by $\text{var}(\boldsymbol{\beta}, \mathbf{b}, H \mid \tau_s, \tau_h, \mathbf{y}) \approx h(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \hat{H})^{-1}$, where $(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \hat{H})$ is the mode of $p(\boldsymbol{\beta},\ \mathbf{b},\ H \mid \tau_h,\ \tau_s,\ \mathbf{y})$. The normal-case variance inflation factor due to adding spatial effects (9) extends to the present non-Gaussian models as the ratio of $h(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \hat{H})^{-1}_{jj}$ and the approximate posterior variance of $\boldsymbol{\beta}_j$ under the nonspatial model,

$$\text{VIF}_j(\tau_s, \tau_h) = \frac{h(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \hat{H})^{-1}_{jj}}{\{X'W(\boldsymbol{\beta}^*, 0, 0)X\}^{-1}_{jj}}. \tag{23}$$

In the normal case, the variance inflation factor (9) was a function of only $r = \tau_s/\tau_e$ and did not depend on τ_e. In the nonnormal case, the variance inflation factor (23) is a function of both τ_s and τ_h. The spatial model for nonnormal outcomes in (19) adds both spatial and heterogeneity random effects to the simple generalized linear model. If either of these random effects is given a flat prior, that is, if either τ_s or τ_h is zero,

$\boldsymbol{\beta}_j$ will be perfectly correlated with some combination of the random effects and VIF_j will be infinite.

The reduced model in Section 3 diminished the variance inflation of the fixed effects caused by collinearity by deleting the spatial smooth linear combinations of the spatial random effects that were correlated with the fixed effects. Extending this method to the nonnormal case requires removing the combinations of both $\mathbf{b}$ and H that are correlated with $\boldsymbol{\beta}$. $h(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \hat{H})$ resembles the Hessian of the linear mixed model with linear predictor $\hat{W}^{1/2}X\boldsymbol{\beta} + \hat{W}^{1/2}Z\mathbf{b} + \hat{W}^{1/2}H$, where $\mathbf{b} \sim N(0, \tau_s D)$, $H \sim N(0, \tau_h I)$, and $\hat{W} = W(\hat{\boldsymbol{\beta}}, \hat{\mathbf{b}}, \hat{H})$. To remove the collinearity between the fixed effects and the random effects in this model, we could delete the combinations of $\hat{W}^{1/2}Z\mathbf{b}$ and $\hat{W}^{1/2}H$ that are in the span of $\hat{W}^{1/2}X$. Following the steps of Section 3 to remove the collinear combinations of $\hat{W}^{1/2}Z\mathbf{b}$ and $\hat{W}^{1/2}H$ gives a reduced generalized linear mixed model analogous to the reduced Gaussian model in (16) and (17), that is,

$$\eta = X\boldsymbol{\beta} + \hat{W}^{-1/2}L'\boldsymbol{\theta}_2 + \hat{W}^{-1/2}L'\boldsymbol{\gamma}_2 \tag{24}$$

$$\boldsymbol{\theta}_2 = L\hat{W}^{1/2}Z\mathbf{b} \sim N\left(0, \tau_s L\hat{W}^{1/2}Q\hat{W}^{1/2}L'\right) \tag{25}$$

$$\boldsymbol{\gamma}_2 = L\hat{W}^{1/2}H \sim N(0, \tau_h L\hat{W}L'), \tag{26}$$

where $\tau_s L\hat{W}^{1/2}Q\hat{W}^{1/2}L'$ and $\tau_h L\hat{W}L'$ are precisions, and L is the $(n - p) \times p$ matrix whose rows are the eigenvectors of $I - \hat{W}^{1/2}X(X'\hat{W}X)^{-1}X'\hat{W}^{1/2}$ that correspond to nonzero eigenvalues.

Let $(\hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\gamma}}_2)$ be the mode of $p(\boldsymbol{\beta},\ \boldsymbol{\theta}_2,\ \boldsymbol{\gamma}_2 \,|\, \tau_h,\ \tau_s,\ \mathbf{y})$ and $\hat{W}_2 = W(\hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\gamma}}_2)$. Using Fisher's observed information approximation, under (24)–(26) $\text{var}(\boldsymbol{\beta},\ \boldsymbol{\theta}_2,\ \boldsymbol{\gamma}_2 \,|\, \tau_h,\ \tau_s,\ \mathbf{y})$ is approximately

$$\begin{pmatrix} X'\hat{W}_2X & X'\hat{W}_2\hat{W}^{-1/2}L' & X'\hat{W}_2\hat{W}^{-1/2}L' \\ L\hat{W}^{-1/2}\hat{W}_2X & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' + \tau_s L\hat{W}^{1/2}Q\hat{W}^{1/2}L' & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' \\ L\hat{W}^{-1/2}\hat{W}_2X & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' & L\hat{W}^{-1/2}\hat{W}_2\hat{W}^{-1/2}L' + \tau_h L\hat{W}L' \end{pmatrix}^{-1}. \tag{27}$$

The posterior of η should be similar under the full model (19) and the reduced model (24)–(26) because the only difference in these models is that redundancies in η have been removed. Because W is a function of η, the estimate of W from the full model ($\hat{W}$) and reduced model ($\hat{W}_2$) should be similar. If $\hat{W} \approx \hat{W}_2$, $X'\hat{W}_2\hat{W}^{-1/2}L' \approx 0$ (because L is, by construction, orthogonal to $\hat{W}^{1/2}X$) and $\boldsymbol{\beta}$ is approximately uncorrelated with $\boldsymbol{\theta}_2$ and $\boldsymbol{\gamma}_2$.

5. Analysis of Slovenia Data

Section 2 predicts that the change in the fixed effects due to adding the CAR parameters will be large if there is little spatial smoothing of the CAR parameters and if the covariate is highly correlated with low-frequency eigenvectors of Q. Both of these criteria are met in the Slovenia analysis: the effective number of parameters in the model is $p_D = 62.3$ (Table 1) out of a possible $n = 192$ and the correlation between SEc and the eigenvector of Q corresponding to the smallest eigenvalue is 0.72. The posterior variance of β_{SEc} increases from 0.0004 under the simple Poisson regression model to 0.0016 under the spatial Poisson regression model.

In this section, the methods of Section 4 are used to investigate influence and collinearity in the analysis of the Slovenia cancer data. These diagnostics rely on a one-step approximation to β_{SEc}'s posterior. For the Slovenia data, this approximation works fairly well. The estimate of β_{SEc}'s variance given by assuming $(\tau_s, \tau_h) = (10.5, 125.9)$, their posterior medians under Model 2, and by inverting the Hessian matrix (21) is 0.0013, which is similar to β_{SEc}'s posterior variance under the spatial Poisson regression model, 0.0016. Figure 3a shows that the one-step estimate for β_{SEc} in (22) (solid line) is near β_{SEc}'s true posterior median (centers of the boxplots) under Model 2 for all values of τ_s that have appreciable posterior mass.

As mentioned in Section 1, the posterior median of β_{SEc} shifts from -0.137 to -0.022 after adding the spatial terms to the simple Poisson regression model. Figure 3a depicts the joint posterior of $(\beta_{\text{SEc}}, \tau_s)$ and shows how this shift in β_{SEc}'s median depends on τ_s, the parameter that controls smoothing of the CAR parameters. For small τ_s, the CAR parameters are unsmoothed and β_{SEc}'s median is close to zero. As τ_s increases, β_{SEc}'s median gradually tends toward β_{SEc}'s posterior median from the simple Poisson regression model (the middle horizontal dashed line). For the range of τ_s with appreciable posterior mass, β_{SEc}'s posterior median does not reach its posterior median from the simple Poisson regression model. The model in (24)–(26) resolves the collinearity problem by removing the combinations of the CAR random effects that are collinear with the fixed effects. Figure 3b shows that deleting the collinear spatial terms removes β_{SEc}'s dependence on τ_s; β_{SEc}'s median under the reduced model is similar for all τ_s and much closer to its posterior median in the nonspatial Poisson model.

The δ_{ij} statistics of Section 4 can be used to measure each region's contribution to the shift in β's posterior due to adding the spatial random effects. These statistics are plotted in Figure 4 for the Slovenia data with (τ_s, τ_h) equal to their posterior medians under Model 2 (Table 1); $\delta_i < 0.010$ for each municipality except the northeastern regions Murska Sobota and Ptuj that have $\delta_i > 0.015$. Murska Sobota and Ptuj are unusual because despite having higher SEc than their neighbors (Figure 1b), they do not have lower SIR than their neighbors (Figure 1a). Comparing these regions to their neighbors contradicts the general pattern that regions with high SEc have low SIR. Removing these municipalities from Model 2 causes β_{SEc}'s posterior median to shift from -0.022 to -0.052, a change that is very close to the sum of the δ_i for Murska Sobota and Ptuj.

Another feature that distinguishes Murska Sobota and Ptuj from their neighbors is that they are 2 of the 11 municipalities the Slovenian government defines as urban. Because rural and urban neighbors are probably less similar than neighboring rural regions, smoothing these two types of neighbor pairs equally may not be appropriate. Smoothing all neighbors equally may be more sensible after including an urban indicator as a fixed effect to account for rural/urban differences. The model with a fixed effect for urban/rural is summarized

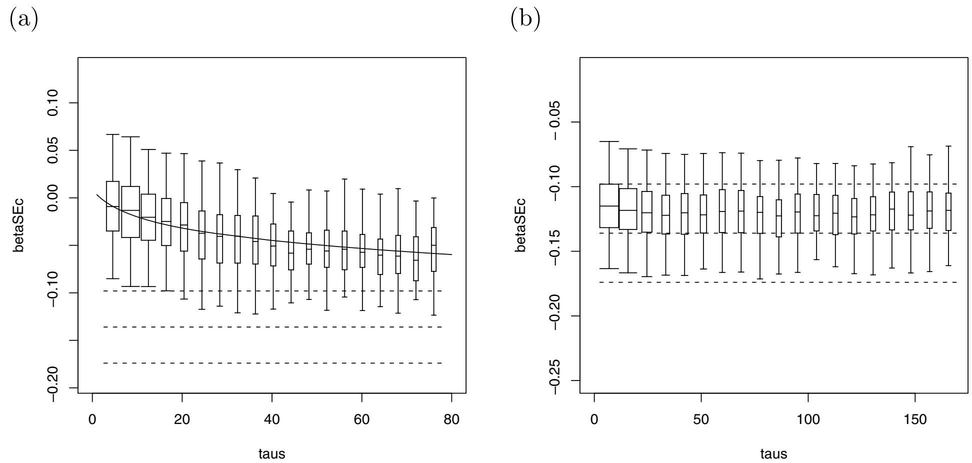

Figure 3. Plot of (β_{SEc}, τ_s)'s joint posterior with the combinations of the random effects that are collinear with SEc (a) included as in Section 1 and (b) excluded as in (24)–(26). The 30,000 MCMC samples from β_{SEc}'s posterior are divided into bins according to the corresponding draws of τ_s. Each panel shows box-plots of the samples of β_{SEc} in each bin with the box-plot's width indicating the number of samples in the bin. The solid line in (a) represents the one-step estimate of β_{SEc} in (22) evaluated at τ_h's posterior median, $\tau_h = 125.9$, and the dashed lines represent β_{SEc}'s posterior median and 95% confidence interval from the Poisson regression without spatial terms.

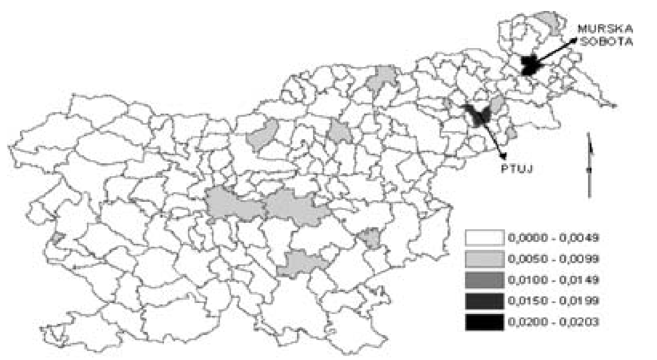

Figure 4. Plot of $|\delta|$, that is, each region's contribution to the change in β's posterior mode due to the spatial terms, with (τ_s, τ_h) set to their posterior medians (10.5, 125.9).

in the third row of Table 1. Adding an urban fixed effect leads to only a small change in β_{SEc}. Thus, urban/rural differences do not mediate the effect of collinearity on β_{SEc}.

The model in (24)–(26) resolves the collinearity problem by removing the combinations of the CAR random effects that are collinear with the fixed effects. Table 1 shows that the posterior of β_{SEc} after removing the collinear spatial effects (Model 4) is indeed more similar to its posterior under the simple Poisson model (Model 1) than its posterior under the usual spatial regression (Model 2). Also, β_{SEc}'s posterior variance is twice as large under Model 4 (0.0008) compared to Model 1 (0.0004). Two factors inflate β_{SEc}'s marginal posterior variance from Model 1 to Model 2: collinearity with the spatial random effects and an increase in the effective number of parameters in the model. The variance of β_{SEc} under Model 4 is smaller than its variance under Model 2 because the collinearity effect is removed but is larger than its variance under Model 1 because the spatial correlation in the data is taken into account.

However, β_{SEc}'s posterior is not identical under Models 1 and 4. Under Model 4, β_{SEc}'s posterior median is -0.120 compared to -0.137 under Model 1. In deriving Model 4 (Section 4), we assumed the posterior modes of e^{η_j} were the same under Models 2 and 4. The slight difference in β_{SEc}'s posterior median under Models 1 and 4 may be due to slight differences in posteriors of e^{η_j} under Models 2 and 4 (Figure 5).

Figure 5a plots the posterior mean relative risks (e^{η_j}) under the usual CAR model (Model 2) and the model without the collinear CAR parameters (Model 4). Although removing the collinear terms from the spatial regression impacts β_{SEc}'s posterior, the posterior mean relative risks under the reduced model are similar to the full spatial model. The remaining spatial random effects in Model 4 are smoothed more than the spatial random effects in Model 2 (the posterior median of τ_s increases from 10.5 under Model 2 to 15.0 under Model 4). However, the remaining heterogeneity random effects in Model 4 are smoothed less than the heterogeneity random effects in Model 2 (the posterior median of τ_h decreases from 125.9 under Model 2 to 38.9 under Model 4). As a result, the fit under the model without the collinear spatial terms is less smooth and neighboring relative risk estimates are less similar (Figure 5b).

The model without the collinear random effects (DIC = 1088.0) has slightly larger DIC than the usual spatial regression (DIC = 1081.5). The models are similar in terms of fit (the posterior mean deviance, $\bar{D}$, is 1019.2 for Model 2 and 1018.0 for Model 4) but removing the collinear random effects gives 7.7 additional effective parameters (the effective model size, p_D, is 62.3 for Model 2 and 70.0 for Model 4), mostly heterogeneity random effects. A researcher focused on estimating each municipality's relative risk would prefer the usual spatial regression because of its smaller DIC. However,

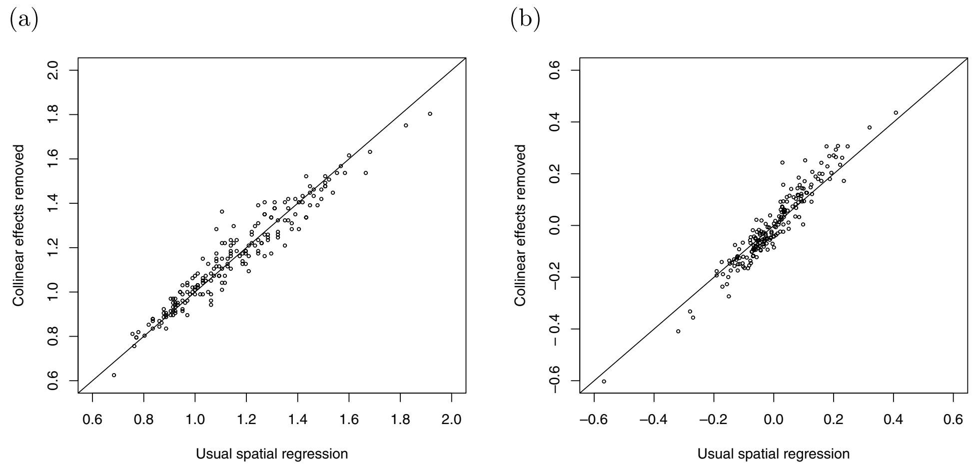

Figure 5. Comparison of the fitted values under the usual spatial regression model of Section 1 and the spatial regression model of Section 4 that excludes the CAR parameters that are collinear with SEc. (a) plots each municipality's posterior mean relative risk (e^{η_j}) and (b) plots $e^{\eta_j} - e^{\eta}_{(j)}$, where $e^{\eta}_{(j)}$ is the average posterior mean relative risk of the municipality j's neighbors.

a researcher primarily focused on measuring the association between socioeconomic status and stomach cancer may prefer the model without the collinear random effects because of the interpretability of the β_{SEc} parameter under this model.

6. Discussion

This article investigated the effect of adding spatial random effects on the posterior distribution of fixed effects in CAR models for disease mapping. Assuming normal outcomes, Section 2.1 proposed the C_{ij} and δ_{ij} statistics to investigate each region's contribution to the change in the fixed effect estimate by including spatial random effects. The Slovenia example illustrated that regions were highly influential if they have different covariates from nearby regions and also if they have large OLS residuals. Section 2.2 showed that the variance inflation from collinearity with the spatial random effects is most troublesome if there is little spatial smoothing and the covariates are highly correlated with low-frequency eigenvectors of the adjacency matrix Q. Sections 3 and 4 developed a model removing the collinear spatial terms and Section 5 applied it to the Slovenia data.

This article's results were developed for the CAR model of areal data but can be extended to other spatial models. A common geostatistical model for point-referenced data is

$$\begin{aligned} \mathbf{y} \mid \boldsymbol{\beta}, S, \tau_e &\sim N(X\boldsymbol{\beta} + S, \tau_e I_n) \\ S \mid \tau_s &\sim N(0, \tau_s Q(\phi)), \end{aligned} \tag{28}$$

where ϕ is a vector of unknown parameters. Mechanically, the only difference between this model and the CAR model in (2) and (3) is that $Q(\phi)$ depends on ϕ. This complicates matters because the geometry of the model, that is, the eigenvectors and eigenvalues of S's prior precision, depends on ϕ. For example, the collinearity between the fixed effects and the random effects was illuminated in (4) and (5) by a transformation that depended on Q's eigenvectors. This transformation would be more difficult to interpret if the eigenvectors themselves were functions of unknown parameters as in (28). However, because the methods developed in Sections 2–4 conditioned on the precisions, (τ_e, τ_s), these methods could be applied to the geostatistical model by simply conditioning on (τ_e, τ_s, ϕ).

Acknowledgement

The authors thank Professor Bradley Carlin of the University of Minnesota for his thoughtful discussion.

References

Besag, J., York, J. C., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* **43,** 1–59.

Christensen, R., Johnson, W., and Pearson, L. M. (1992). Prediction diagnostics for spatial linear models. *Biometrika* **79,** 583–591.

Christensen, R., Pearson, L. M., and Johnson, W. (1993). Case-deletion diagnostics for mixed models. *Technometrics* **34,** 133–169.

Clayton, D. G. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrika* **43,** 671–681.

Haslett, J. (1999). A simple derivation of deletion diagnostic results for the general linear model with correlated errors. *Journal of the Royal Statistical Society, Series B* **61,** 603–609.

Hocking, R. R. (1996). *Methods and Applications of Linear Models: Regression and the Analysis of Variance.* New York: Wiley Interscience.

Hodges, J. S., Carlin, B. P., and Fan, Q. (2003). On the precision of the conditionally autoregressive prior in spatial models. *Biometrics* **59,** 317–322.

Lee, Y. and Nelder, J. A. (1996). Hierarchical generalized linear models. *Journal of the Royal Statistical Society Series B* **58,** 619–678.

Mackinnon, M. J. and Puterman, M. L. (1998). Collinearity in generalized linear models. *Communications in Statistics—Theory and Methods* **18,** 3463–3472.

Reich, B. J. and Hodges, J. S. (2005). *Identification of the variance components in the general two-variance linear model*. Research Report 2005-009, Division of Biostatistics, University of Minnesota, Minneapolis. Available at `http://www.biostat.umn.edu/rrs.php`.

Schaefer, R. L. (1979). Multicollinearity in logistic regression. Ph.D. Thesis, University of Michigan, Ann Arbor.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion and rejoinder). *Journal of the Royal Statistical Society, Series B* **64,** 583–639.

Zadnik, V. and Reich, B. J. (2006). Analysis of the relationship between socioeconomic factors and stomach cancer incidence in Slovenia. *Neoplasma* **53,** 103–110.

Received March 2006. *Revised May* 2005.
Accepted March 2005.